

Analisis Sentimen Tingkat Kepuasan Pelanggan Terhadap Layanan Kurir *J&T Express* Menggunakan Algoritma *Support Vector Machine (SVM)* Berdasarkan Ulasan Pengguna Di *Google Playstore*

Indar Swarsih¹, Suzuki Sofian², Joko Tri Handoko³, MS Hasibuan⁴

¹ Magister Manajemen Teknologi, IIB Darmajaya, ² Magister Manajemen Teknologi, IIB Darmajaya

³ Magister Manajemen Teknologi, IIB Darmajaya

indar.2222320025@gmail.darmajaya.ac.id
suzuki.2222320016@gmail.darmajaya.ac.id
joko.2222320010@mail.darmajaya.ac.id
msaid@darmajaya.ac.id

Abstract — Courier service is one of the services that are widely used by the public, especially in the current digital era. In this context, courier services allow shippers to deliver goods or documents without the need to be present directly to the destination location. J & t express, as one of the shipping expedition service providers in Indonesia, is the first choice for many people. Although technology continues to evolve and competition is increasingly fierce, the quality of courier services is a key factor that customers need to pay attention to. However, it should be noted that the J&t express application in the google play store received a low rating, and this is the background of this study.

The main focus of this study was to identify the level of customer satisfaction with J&t express courier services through reviews available on the google play store. Within the framework of this study, sentiment analysis was conducted using the support vector machine algorithm, by applying crisp-dm methodology. The results showed that from the business understanding stage to the modeling stage, the performance of the support vector machine can be considered good. In addition, this study also resulted in an implementation that can be accessed through a website with the address jnt-sentiment.streamlit.app. Hopefully, this research can contribute to j & t express in understanding the views of customers and improving the quality of their services

Keyword — analisis sentimen, crisp-dm, support vector machine, j&t express, google play store.

Keyword — Delay response, dual-band operation, filter, LTE, microstrip line.

ABSTRAK — Layanan kurir merupakan salah satu layanan yang banyak dimanfaatkan oleh masyarakat, khususnya di era digital saat ini. Dalam konteks ini, layanan kurir memungkinkan pengirim untuk mengirimkan barang atau dokumen tanpa perlu hadir langsung ke lokasi tujuan. J&T Express, sebagai salah satu penyedia layanan ekspedisi pengiriman di Indonesia, menjadi pilihan utama bagi banyak orang. Walaupun teknologi terus berkembang dan persaingan semakin sengit, kualitas layanan kurir menjadi faktor kunci yang perlu diperhatikan oleh pelanggan. Namun, perlu dicatat bahwa aplikasi J&T Express di Google Play Store mendapatkan rating yang rendah, dan hal ini menjadi latar belakang dari penelitian ini.

Fokus utama dari penelitian ini adalah untuk mengidentifikasi tingkat kepuasan pelanggan terhadap layanan kurir J&T Express melalui ulasan yang tersedia di Google Play Store. Dalam kerangka penelitian ini, dilakukan analisis sentimen menggunakan algoritma Support Vector Machine, dengan menerapkan metodologi CRISP-DM. Hasil penelitian menunjukkan bahwa dari tahap pemahaman bisnis hingga tahap pemodelan, performa Support Vector Machine dapat dianggap baik. Sebagai tambahan, penelitian ini juga menghasilkan implementasi yang dapat diakses melalui situs web dengan alamat jnt-sentiment.streamlit.app. Harapannya, penelitian ini dapat memberikan kontribusi bagi J&T Express dalam memahami pandangan pelanggan dan meningkatkan kualitas layanan mereka.

Kata kunci — Analisis Sentimen, Crisp-Dm, Support Vector Machine, J&T Express, Google Play Store

I. PENDAHULUAN

1.1 Latar Belakang Masalah

Perkembangan teknologi informasi dalam beberapa dekade terakhir telah memungkinkan pengolahan data yang lebih canggih dan efisien. Dalam konteks ini, teknologi data mining menjadi sangat vital dalam membantu organisasi mendapatkan informasi yang lebih mendalam dan akurat dari data yang tersimpan dalam basis data. Dengan kemampuan untuk mengolah data yang lebih besar dan lebih kompleks, teknologi data mining memainkan peran yang semakin krusial dalam proses pengambilan keputusan.

Analisis sentimen merupakan bagian dari Pemrosesan Bahasa Alami (Natural Language Processing/NLP) yang mengembangkan sistem untuk mengidentifikasi dan mengekstrak pendapat dalam bentuk teks. Saat ini, informasi berupa teks banyak tersedia di internet dalam format forum, blog, media sosial, dan situs yang berisi ulasan. Dengan bantuan analisis sentimen, informasi yang semula tidak terstruktur dapat diubah menjadi data yang lebih terstruktur.

Pelayanan kurir merupakan salah satu jenis layanan pengiriman barang atau dokumen yang banyak dimanfaatkan masyarakat, terutama di era digital ini. Dengan adanya layanan kurir, pengirim tidak perlu datang langsung ke tujuan untuk mengirimkan barang atau dokumen, melainkan dapat memanfaatkan jasa pengiriman yang telah tersedia. Di antara banyaknya ekspedisi pengiriman di Indonesia, salah satunya adalah J&T Express. J&T Express menyediakan layanan pengiriman yang mencakup berbagai jenis barang atau produk ke seluruh wilayah Indonesia, memberikan kemudahan bagi masyarakat dalam memenuhi kebutuhan mereka.

Meski demikian, dengan kemajuan teknologi dan persaingan yang semakin intensif, kualitas layanan kurir menjadi aspek yang krusial bagi para pelanggan. Keterlambatan pengiriman, kerusakan barang, dan pelayanan yang kurang memuaskan dapat memberikan dampak negatif terhadap tingkat kepercayaan dan reputasi dari jasa kurir tersebut.

Aplikasi J&T Express di Google Play Store memiliki rating yang rendah, itulah mengapa penelitian ini dilakukan. Fokus dari penelitian ini adalah mengeksplorasi tingkat kepuasan terhadap pelayanan kurir yang tercermin dalam ulasan Google Play Store, dengan tujuan mengidentifikasi permasalahan yang mungkin terjadi pada layanan kurir J&T Express. Selain itu, penelitian ini akan melibatkan analisis sentimen menggunakan algoritma Support Vector Machine (SVM) untuk mengukur kepuasan pelanggan terhadap layanan kurir yang diberikan oleh J&T Express. Sebagai sumber data, penelitian ini menggunakan hasil web scraping dari Google Play Store untuk mengambil ulasan dari pengguna aplikasi J&T Express. Harapannya, penelitian ini dapat memberikan kontribusi bagi J&T Express dalam memahami perspektif pelanggan dan meningkatkan kualitas layanan yang mereka berikan.

Berdasarkan uraian di atas, berikut adalah rumusan masalah penelitian ini.

1. Berapakah tingkat akurasi yang dihasilkan oleh metode SVM dalam melakukan klasifikasi tingkat kepuasan pelanggan di J&T Express?
2. Bagaimana kinerja metode *Support Vector Machine* (SVM) dalam klasifikasi tingkat kepuasan pelanggan di J&T Express?

Tujuan penelitian ini untuk mengetahui tingkat kepuasan pelanggan terhadap layanan kurir J&T Express dengan menggunakan algoritma SVM, mengetahui akurasi dari algoritma SVM, mengetahui performa algoritma SVM, serta menerapkan analisis sentimen J&T Express pada sebuah web dengan model yang sudah dilatih. Manfaatnya dari penelitian ini dapat memberikan informasi kepada pembaca bagaimana tingkat kepuasan pelanggan terhadap layanan kurir J&T Express.

2. LANDASAN TEORI

2.1 Analisis Sentiment

Analisis sentimen, menurut penelitian yang dilakukan oleh Pasek et al. pada tahun 2022, adalah proses mengklasifikasikan emosi (positif, negatif, dan netral) yang ditemukan dalam tulisan dengan menggunakan metode analisis. Analisis sentimen juga disebut Opinion Mining, yang merujuk pada bidang yang luas dengan tujuan melakukan analisis terhadap pendapat, evaluasi sentimen, sikap, penilaian, dan emosi seseorang. Dari pernyataan ini, dapat disimpulkan bahwa analisis sentimen adalah proses penentuan perasaan atau tanggapan seseorang, yang dapat diungkapkan melalui bahasa atau teks dan dikategorikan sebagai sentimen netral, positif, atau negatif.

2.2 Data Mining

Berdasarkan penelitian yang dilakukan oleh Pratama dkk. pada tahun 2022, Data Mining adalah suatu proses penemuan korelasi, pola, dan tren yang bermanfaat dengan mengeksplorasi sejumlah besar data dari berbagai repository, menggunakan teknologi pengenalan pola seperti statistik dan teknik matematika. Data Mining merupakan kegiatan analisis data yang melibatkan penggunaan perangkat lunak untuk mengidentifikasi pola dan aturan dalam kumpulan data. Metode ini digunakan untuk menemukan pengetahuan yang terkandung dalam suatu set data yang berskala besar.

2.3 Text Mining

Menurut penelitian yang dilakukan oleh Ramadhani & Wahyudin pada tahun 2022, Text Mining adalah suatu proses dalam domain Data Mining yang didefinisikan sebagai penambangan data dalam teks atau data yang biasanya ada dalam dokumen. Tujuan Text Mining adalah untuk menemukan informasi baru atau frasa tertentu, dan untuk melakukan proses ini diperlukan analisis yang saling terkait antar dokumen.

2.4 CRISP-DM

Menurut penelitian yang dilakukan oleh Kurniawan & Yasir pada tahun 2022, Cross Industry Standard Process for Data Mining (CRISP-DM) adalah suatu model proses yang independen dari industri yang umumnya digunakan untuk keperluan Data Mining. Metode ini terdiri dari enam fase yang berulang. Dimulai dari fase Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation, dan Deployment. Dalam pengembangannya, CRISP-DM dianggap sebagai metodologi Data Mining yang paling komprehensif dalam memenuhi kebutuhan proyek industri, dan telah menjadi yang paling banyak digunakan dalam proyek analitik, Data Mining, serta ilmu Data.

2.5 Preprocessing

Penelitian yang dilakukan oleh Gifari et al. pada tahun 2022 menunjukkan bahwa proses preprocessing berfungsi untuk menormalkan istilah yang berasal dari kalimat. Proses ini dilakukan untuk memastikan bahwa data latih dan fitur yang diekstrak nantinya sesuai dengan data yang diinginkan. Hal

ini bertujuan untuk membuat proses pengolahan data lebih mudah. Enam proses terlibat dalam pra-pengolahan, sebagai berikut :

1. Normalisasi: Tahap ini dilakukan untuk memastikan bahwa kata-kata yang diperpanjang atau disingkat menjadi kata-kata yang normal menurut Kamus Besar Bahasa Indonesia (KBBI). Mengubah kata tidak baku menjadi kata baku disebut "konversi slang".
2. Pembersihan: Pada tahap ini, aksara non-alfabetis dihapus, seperti titik (.), koma (,), tanda tanya (?), dan tanda seru (!), serta simbol-simbol seperti tanda "@" untuk nama pengguna, hashtag (#), emoticon, dan alamat web.
3. Folding Cases: Tahap ini mengubah karakter alfabet yang telah dibersihkan ke huruf kecil, juga dikenal sebagai folding case.
4. Tokenizing: Fase ini memecah kalimat berdasarkan kata-kata yang membentuknya, yang disebut term atau token. Ini dipecah berdasarkan spasi.
5. *Filtering*: Untuk analisis sentimen, kata-kata yang tidak penting atau tidak bermakna dihapus. Ada kata-kata seperti atau, yang, dengan, di, ke, dan tetapi.
6. Stemming: Tahap ini berfungsi untuk mengubah kata berimbuhan menjadi kata dasar.

2.6 Support Vector Machine

Menurut penelitian yang dilakukan oleh Suryati dkk. pada tahun 2023, Support Vector Machine (SVM) adalah salah satu algoritma machine learning yang sangat terkenal untuk keperluan klasifikasi. Dalam dekade terakhir, SVM telah menjadi metode yang kuat untuk pola klasifikasi dan telah mencapai tingkat keberhasilan yang tinggi di berbagai bidang. Oleh karena itu, banyak dari komunitas machine learning yang tertarik untuk mempelajari dan mengembangkan SVM karena kinerjanya yang sangat baik dalam menangani berbagai masalah pembelajaran. SVM merupakan metode machine learning yang bertujuan untuk menemukan hyperplane terbaik yang dapat memisahkan dua kelas di dalam ruang input (input space). Algoritma klasifikasi SVM menggunakan data pelatihan untuk membentuk model klasifikasi, dan model yang terbentuk digunakan untuk memprediksi kelas data baru yang belum pernah ada sebelumnya, yang disebut sebagai data pengujian.

3. METODE PENELITIAN

3.1 Pengumpulan Data

Metode pengumpulan data yang diterapkan dalam penelitian "Analisis Sentimen Tingkat Kepuasan Pelanggan Terhadap Layanan Kurir J&T Express Menggunakan Algoritma Support Vector Machine (SVM) Berdasarkan Review Pengguna Di Google Playstore" mencakup:

1. Studi Literatur: Peneliti mengumpulkan Data dari jurnal atau karya tulis ilmiah yang relevan, sehingga dapat membantu peneliti dalam

menambah referensi sesuai dengan topik yang diangkat.

2. Filter Data: Peneliti menggunakan filtering Data dari Dataset yang akan digunakan, untuk mendapatkan variabel dan Data yang dapat digunakan.

3.2 Metode Implementasi

Penerapan model yang dijelaskan oleh peneliti menggunakan metode CRISP-DM (Cross Industry Standard Process for Data Mining) melibatkan beberapa tahap, sebagai berikut:

1. Pemahaman Bisnis (Business Understanding): Pada tahap ini, penelitian memahami dan mengumpulkan data melalui metode Text Mining di Google Play Store.
2. Pemahaman Data (Data Understanding): Data dianalisis untuk memastikan apakah sudah memadai dan layak untuk diproses atau apakah perlu dilakukan pengumpulan data tambahan.
3. Persiapan Data (Data Preparation): Melalui proses Text Preprocessing, data dibersihkan dan diolah sehingga menghasilkan data yang berkualitas. Selanjutnya, dilakukan labeling atau pelabelan kata-kata sebagai positif dan negatif.
4. Modeling: Pada tahap ini, dilakukan pembobotan kata positif dan negatif menggunakan TF-IDF. Algoritma Support Vector Machine kemudian diterapkan pada data latih dan data uji untuk mendapatkan tingkat akurasi.
5. Evaluasi (Evaluation): Setelah akurasi diperoleh, algoritma Support Vector Machine yang telah dilatih dan diuji dievaluasi menggunakan Confusion Matrix untuk menilai sejauh mana kemampuannya dalam mengklasifikasi data. Hasil model yang sudah dilatih dan diuji dapat disimpan menggunakan Python pickle.
6. Implementasi (Deployment): Selanjutnya, untuk dapat diakses oleh pengguna, Analisis Sentimen akan di-deploy atau di-hosting menggunakan web framework Streamlit, dan Deta Space sebagai Cloud Database

4. PEMBAHASAN

4.1 Business Understanding

Tujuan dari penelitian ini adalah melakukan analisis terhadap ulasan pengguna terkait aplikasi J&T Express dengan maksud untuk menentukan apakah ulasan tersebut cenderung bersifat positif atau negatif. Langkah ini diambil agar J&T Express dapat memahami permasalahan yang mungkin muncul dalam layanannya. Dengan mengetahui permasalahan tersebut, J&T Express berharap dapat meningkatkan kualitas pelayanannya guna memberikan pengalaman yang lebih baik. Rating yang rendah untuk aplikasi J&T Express (2,4 dari 5 bintang) menunjukkan adanya keluhan atau ketidakpuasan dari pengguna, sehingga menjadi sangat penting bagi J&T Express untuk mengidentifikasi dan memahami permasalahan yang muncul.

Dalam kerangka penelitian ini, peneliti mengumpulkan dataset ulasan pengguna J&T Express dan menjalankan

analisis menggunakan metode Support Vector Machine untuk mengklasifikasikan apakah ulasan tersebut bersifat positif atau negatif. Penelitian ini juga bertujuan untuk meraih informasi baru dari dataset ulasan pengguna J&T Express, sekaligus untuk mengevaluasi performa hasil klasifikasi yang diperoleh dengan menggunakan metode Support Vector Machine.

4.2 Data Understanding

Pengumpulan data dilaksanakan melalui metode Pemulasaran Data ulasan pengguna aplikasi J&T Express di Google Play Store, menggunakan perpustakaan (library) yang tersedia pada Python, yaitu google_play_scraper. Jumlah data awal yang diambil adalah sebanyak 9000 data yang dianggap paling relevan. Kemudian, dilakukan proses penyaringan data dengan menghilangkan kata "aplikasi" atau "apk" agar data lebih fokus pada layanan kurir, sehingga jumlah data berkurang menjadi 7815.

4.3 Data Preparation

Sebelum data dimasukkan ke dalam proses pemodelan, proses pembersihan atau penyempurnaan data dilakukan. Langkah ini dilakukan dengan tujuan untuk menghasilkan model dan akurasi yang ideal sambil memfasilitasi peneliti manajemen data. Dalam kasus ini, peneliti memilih empat sampel data untuk diproses pada tahap persiapan data. Salah satu dari empat data dalam bagian "isi" ditunjukkan pada Gambar 1.

content
Pelayanan sangat buruk, barang telah tertunda selama 2 hari, membuat saya merasa kecewa.
Sangat mengecewakan, setelah menunggu lama, paket justru hilang.
Layanan nya murah dan dijamin cepat
saya kecewa dengan pelayanan jnt, buruk!!!

Gambar 1 data sampel pada content

Berikut tahapan pada Data Preparation yang telah peneliti implementasi:

4.3.1 Case Folding

Untuk membuat pengelolaan data lebih mudah, proses Case Folding mengubah teks menjadi huruf kecil dan menghapus karakter tertentu. Sebagai contoh, kata "pelayanan" akan dikecilkan menjadi "pelayanan".

caseFolding
pelayanan sangat buruk, barang telah tertunda selama 2 hari, membuat saya merasa kecewa.
sangat mengecewakan, setelah menunggu lama, paket justru hilang.
layanan nya murah dan dijamin cepat
saya kecewa dengan pelayanan jnt, buruk!!!

Gambar 2 Proses Case Folding

4.3.2 Cleansing

Pada tahap Cleansing dilakukan dengan membersihkan teks dari noise atau elemen yang tidak diperlukan, meliputi

angka, tanda baca, emoji dan spasi berlebih. Sebagai contoh kata "2 hari" angkanya di hilangkan menjadi "hari", dan "jelekkk!!!" tanda serunya dihilangkan menjadi "jelekkk"

cleansing
pelayanan sangat buruk barang telah tertunda selama hari membuat saya merasa kecewa
sangat mengecewakan setelah menunggu lama paket justru hilang
layanan nya murah dan dijamin cepat
saya kecewa dengan pelayanan jnt buruk

Gambar 3 Proses Cleansing

4.3.3 Stemming

Pada tahap stemming dilakukan dengan menghilangkan imbuhan atau akhiran kata dalam teks sehingga hanya menyisakan bentuk dasarnya. Sebagai contoh kata "pelayanan" imbuhnya di hilangkan menjadi "layan".

stemmer
layan sangat buruk barang telah tunda lama hari buat saya rasa kecewa
sangat kecewa telah tunggu lama paket justru hilang
layan nya murah dan jamin cepat
saya kecewa dengan layan jnt buruk

Gambar 4 Proses Stemming

4.3.4 Slang Word Standardization

Pada tahap slang word standardization dilakukan dengan mengubah kata gaul "slang" menjadi kata formal menggunakan kamus slang word yang sudah peneliti buat. Sebagai contoh kata "jelekkk" di ubah kata gaulnya menjadi formal "jelek"

slang
layan sangat buruk barang telah tunda lama hari buat saya rasa kecewa
sangat kecewa telah tunggu lama paket justru hilang
layan nya murah dan jamin cepat
saya kecewa dengan layan jnt buruk

Gambar 5 Proses Slang Word Standardization

4.3.5 Stop Word Removal

Pada tahap stop word removal dilakukan dengan menghapus kata-kata yang umum dan tidak memiliki informasi penting dalam teks. Sebagai contoh kata "saya" di hilangkan.

stopword
layan buruk barang tunda kecewa
kecewa tunggu paket hilang
layan murah jamin cepat
kecewa layan jnt buruk

Gambar 6 Proses Stop Word Removal

4.3.6 Unwanted Word Removal

Pada tahap ini, kata-kata yang telah dibuat oleh peneliti dibuang, termasuk nama bulan seperti "januari" dan "jan" seterusnya.

stopword
layan buruk barang tunda kecewa
kecewa tunggu paket hilang
layan murah jamin cepat
kecewa layan jnt buruk

Gambar 7 Proses Unwanted Word Removal

4.3.7 Short Word Removal

Pada tahap pertama, kata-kata yang di bawah tiga huruf dihilangkan karena tidak memiliki arti yang signifikan.

4.3.8 Tokenizing

Pada tahap tokenizing dilakukan dengan membuat Split Text (memisahkan kata-kata) menjadi per kata untuk masing-masing ulasan.

4.3.9 Labeling Data

Pada Tahap labeling Data dilakukan pelabelan Data agar mengetahui kalimat bersifat positif atau negatif dengan menggunakan kamus positif dan negatif yang sudah dibuat oleh penulis lalu dihitung jumlahnya, jika suatu kalimat mengandung kata positif akan +1 dan jika mengandung kata negatif akan -1.

4.3.10 Term Frequency-Inverse Document Frequency (TF-IDF)

Tahap TF-IDF (Term Frequency-Inverse Document Frequency) dilakukan dengan menghitung bobot nilai setiap kata dalam teks. Bobot ini menunjukkan seberapa sering kata tersebut muncul dalam dokumen tertentu (Term Frequency) dan seberapa umum atau jarang kata tersebut muncul di setiap dokumen.

4.3.11 Modeling

Berikut adalah tahapan dari modeling menggunakan library dan perhitungan manual.

4.3.12 Split Data

Pada tahap modeling menggunakan library, pertama-tama akan dilakukan split Data (pemisahan Data) menjadi Data uji dan Data test pembagian Data ini menjadi 90% untuk Data uji dan 10% untuk Data test.

4.3.13 Penerapan Library Support Vector Machine

Terakhir akan dilakukan tahap pemodelan dengan Support Vector Machine kernel liner untuk melihat akurasi yang di hasilkan oleh Support Vector Machine.

4.3.14 Perhitungan Manual

Pertama-tama, dalam melakukan perhitungan manual, data akan di-split dari empat sampel, terdiri dari tiga data latih dan satu data uji. Kemudian, tf-idf akan diaplikasikan pada data latih, seperti yang terlihat pada gambar berikutnya. Selanjutnya, kita akan menghitung TF (Term Frequency) dan menjumlahkannya untuk setiap dokumen, sebagaimana terlihat dalam gambar berikutnya. Tahap berikutnya melibatkan normalisasi tf menggunakan rumus tertentu. Proses selanjutnya adalah mencari df dengan mengidentifikasi dokumen mana yang mencakup suatu kata. Sebagai contoh, untuk kata "layan", kata tersebut muncul di dalam dua dokumen, sehingga hasil df-nya adalah 2. Langkah terakhir dalam perhitungan adalah menghitung TF-IDF dengan mengalikan tf dengan hasil idf, seperti yang ditunjukkan dalam rumus $tf\ idf = tf \times idf$. Berikut adalah penyelesaian untuk kata "layan" berdasarkan rumus tf-idf yang telah disebutkan sebelumnya, yaitu $tf\ idf = 0,2 \times 0,301 = 0,06$.

Setelah melalui tahap tf-idf, langkah berikutnya adalah proses pemodelan menggunakan Support Vector Machine dengan kernel linear. Diperlukan nilai support vector (a) dan bobot vector (w) untuk mendapatkan hyperplane (b). Berikutnya adalah contoh proses pelatihan pemodelan data menggunakan data latih yang telah diberi bobot sebelumnya. Selanjutnya, hasil pembobotan kata pada data latih diubah menjadi bentuk vektor. Dokumen yang berlabel positif akan dinyatakan dengan angka 1, sementara dokumen yang berlabel negatif akan dinyatakan dengan angka -1. Berikut adalah vektor pada dokumen D1, D2, dan D3.

4.4 Evaluation

Pada tahap evaluasi, dilakukan pengujian kinerja model dengan menggunakan dua proses, yaitu classification report dan confusion matrix. Classification Report adalah sebuah laporan yang memberikan ringkasan tentang performa model dalam memprediksi sentimen dari teks. Laporan ini menggunakan metrik evaluasi seperti presisi (precision), recall, dan f1-score untuk mengukur kualitas prediksi sentimen. Presisi mengukur tingkat akurasi model dalam memprediksi sentimen positif atau negatif. Recall mengukur seberapa baik model dalam menemukan dan mengklasifikasikan dengan benar sentimen positif atau negatif. F1-score adalah skor yang menggabungkan presisi dan recall untuk memberikan penilaian keseluruhan tentang performa model.

Selain itu, *Classification Report* juga memberikan informasi tentang support, yang menunjukkan jumlah sampel yang termasuk dalam setiap kategori sentimen. Selanjutnya, terdapat confusion matrix, yang merupakan tabel yang digunakan dalam analisis sentimen untuk mengukur kualitas prediksi model antara sentimen positif dan negatif. Tabel ini membandingkan hasil prediksi model dengan label sebenarnya dari data yang dianalisis. Confusion Matrix terdiri dari empat bagian, yaitu True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN). True Positive menunjukkan jumlah prediksi yang benar bahwa sentimen adalah positif. True Negative menunjukkan jumlah prediksi yang benar bahwa sentimen

adalah negatif. False Positive adalah jumlah prediksi yang salah bahwa sentimen adalah positif, sementara False Negative menunjukkan jumlah prediksi yang salah bahwa sentimen adalah negatif. Jika jumlah TP dan TN lebih besar dibandingkan FP dan FN, itu menunjukkan bahwa model mampu mengenali sentimen positif dan negatif dengan baik.

4.5 Deployment

Tahap Deployment merupakan langkah akhir dalam metodologi CRISP-DM, di mana proses Data mining diimplementasikan ke dalam sebuah situs web. Peneliti menggunakan Streamlit sebagai *platform web hosting* dan *Data Space* sebagai penyimpanan cloud untuk data.

5. PENUTUP

5.1 Kesimpulan

Hasil penelitian diuraikan secara rinci di bagian diskusi Sebagai berikut:

1. Berdasarkan kinerja metode Support Vector Machine (SVM), SVM mampu mencapai tingkat akurasi sebesar 96,56%. Untuk mempermudah pemahaman, angka tersebut dibulatkan menjadi 97%. Menurut peneliti, performa ini dianggap sudah cukup baik.
2. Berdasarkan analisis sentimen terhadap aplikasi J&T Express, terdapat lebih banyak polaritas negatif daripada positif. Sentimen negatif yang terakumulasi menunjukkan keluhan terkait "pelayanan buruk dan mengecewakan karena lambatnya pengiriman paket, paket salah kirim, dan kehilangan paket." Di sisi lain, sentimen positif yang terkumpul menggambarkan pujian terhadap "pelayanan yang baik, kurir yang bekerja dengan baik karena pengiriman yang cepat, dan barang sesuai."

5.2 Saran

1. Untuk penelitian mendatang, disarankan agar menggunakan dataset yang lebih besar guna mengumpulkan informasi yang lebih meluas dan mencakup variasi yang lebih besar. Selain itu, disarankan juga untuk menggunakan kamus yang terkini agar data tetap relevan, dan mempertimbangkan penggunaan algoritma lain sebagai perbandingan.
2. Bagi J&T Express, disarankan untuk memperhatikan hasil analisis dan implementasi yang telah dilakukan dalam penelitian ini. Diharapkan perusahaan dapat melakukan perbaikan pada layanannya berdasarkan sentimen dari ulasan pengguna, sehingga J&T Express dapat mengalami perkembangan yang lebih baik ke depannya.

DAFTAR PUSTAKA

- [1] Aziz, A. (2022). Analisis Sentimen Identifikasi Opini Terhadap Produk, Layanan dan Kebijakan Perusahaan Menggunakan Algoritma TF-IDF dan SentiStrength. Dalam Jurnal Sains Komputer & Informatika (J-SAKTI (Vol. 6, Nomor 1).
- [2] Deta Base – Space Docs (2023), Diakses pada 19 Juni 2023, dari <https://deta.space/docs/en/build/reference/deta-base>
- [3] Getting Started – Visual Studio Code Docs (2023), diakses pada 19 juni 2023, dari <https://code.visualstudio.com/docs>
- [4] Gifari, O. I., Adha, M., Rifky Hendrawan, I., Freddy, F., & Durrand, S. (2022). Analisis Sentimen Review Film Menggunakan TF-IDF dan *Support Vector Machine*. Jifotech (Journal Of Information Technology, 2(1).
- [5] Hasna, S. K. (2021). Analisis Sentimen *Data Ulasan Menggunakan Algoritma Support Vector Machine*. Tugas Akhir, Universitas Islam Indonesia
- [6] Kurniawan, D., & Yasir, D. M. (2022). Optimization Sentiment Analysis Using Crisp-Dm And Naïve Bayes Methods Implemented On Social Media. 6, 74–84.
- [7] Locarso, G. K. (2022). Analisis Sentimen Review Aplikasi Pedulilindungi Pada *Google Play Store Menggunakan Nbc*. Jurnal Teknik Informatika Kaputama (JTIK), 6(2).
- [8] Pasek, P., Mahawardana, O., Sasmita, G. A., Agus, P., & Pratama, E. (2022). Analisis Sentimen Berdasarkan Opini dari Media Sosial Twitter terhadap “Figure Pemimpin” Menggunakan Python. Dalam JITTER-Jurnal Ilmiah Teknologi dan Komputer (Vol. 3, Nomor 1).
- [9] Pratama, A. E., Ariesta, A., & Gata, G. (2022). Analisis Sentimen Masyarakat terhadap Tim Nasional Indonesia pada Piala AFF 2020 Menggunakan Algoritma K-Nearest Neighbors The researcher uses the Cross-Industry Standard Process for *Data Mining (CRISP-DM) method and implements the K-Nearest*. Jurnal TICOM: Technology of Information and Communication, 10(3), 187–196.
- [10] Ramadhani, S. H., & Wahyudin, M. I. (2022). Analisis Sentimen Terhadap Vaksinasi Astra Zeneca pada Twitter Menggunakan Metode Naïve Bayes dan K-NN. Jurnal

Teknologi Informasi dan Komunikasi),
6(4), 2022.

Algoritma *Support Vector Machine*
(SVM). 4(1), 96–106.

[11] Suryati, E., Ari Aldino, A., Penulis
Korespondensi, N., & Suryati Submitted,
E. (2023). Analisis Sentimen Transportasi
Online Menggunakan Ekstraksi Fitur
Model Word2vec Text Embedding Dan

[12] *Streamlit* Documentation – *Streamlit* Docs
(2023), Diakses pada 19 Juni 2023, dari
<https://docs.streamlit.io/>